A Formalization of Descartes' Causal Argument for Existence of God

Vincenzo Ciccarelli

Universidade Federal do Rio Grande do Norte, Brasil
`ciccarelli.vin@gmail.com`

Abstract.   In this paper I attempt a formalization of Descartes' causal argument for the existence of God presented in the Third Meditation. By taking into account the philosophical setting of Descartes' Meditations, I argue that the best logical theory to carry out the formalization is quantified modal logic where free logic is used for the first-order axioms. This fact is related to Descartes' distinction between formal and objective reality which — to my mind — suggests both the adoption of modality (e.g. true in formal vs true in objective reality) and the possibility of using empty names (i.e. names which lack reference in "formal reality"). After presenting the formal system and the reconstruction of Descartes' proof, I will discuss some issues related to the philosophical import of the formalization, such as the understanding of rigid designation and lack of reference. I will argue that the formalization of Descartes' argument has an intrinsic philosophical value for it sheds light on the structure of argument, its implicit difficulties, and the use of logic itself to understand the philosophical tradition.

Keywords: Descartes • meditations • causal argument • quantified modal logic • free logic

1. Introduction

In the Third Meditation Descartes (1984) proposes a quite original and complex argument for the existence of God. The argument is commonly labeled as 'causal', insofar as it is in line with the traditional Scholastic proofs of the existence of the cause (the Divine Necessary Being) from a certain effect which is assumed as given. However, Descartes' argument represents a novelty within this tradition in what he considers to be the 'effect': it is not an argument which infers the existence of a creator or architect from the existence of the universe with its structure, for that would be a far too demanding assumption from the perspective of the methodic doubt; yet the existence of God is proven from the fact that exists the idea of an infinite and perfect being in my mind. Perhaps we should call it *an idealist argument*, at least in order to distinguish it within the varieties of causal arguments of Medieval and Modern Metaphysics.

The existence of God is a crucial point in the development of Descartes' epistemic project. For only the existence of a perfect, infinitely benevolent Being may

rule out the hypothesis of the deceiving daemon, and thus ensure the veracity of all ideas which I perceive in a clear and distinct way. Therefore, in the Third Meditation Descartes has to take a break in what has seemed to be thus far a mere epistemological analysis and do some "heavy Metaphysics". What is epistemically available at the outset seems to be a very poor set of premises: the existence of my mind with all of its mental states and representations. Such a difficult situation forces Descartes to find within his mind a special content that allows him to construct a causal argument.

In this paper I will attempt a formalization of Descartes' causal argument using a modal logic extended to the first-order using the axioms of free logic. The task is challenging inasmuch as our modern day logical systems are not always compatible with some fundamental ideas about existence and quantification very common before the XIX century. However, I believe that the juice is worth the squeeze. As philosophy students, we are taught in our logic classes (introductory or advanced) that formal tools are interestingly useful to analyze philosophical arguments. Such a statement probably sounds a bit abstract and distant from the actual practice of studying the most diverse styles of philosophical texts. There are several reasons for which we should attempt the understanding of the logical structure of philosophical arguments using a certain deal of technicality. Firstly, we must be in position to assess the validity of the arguments, a notion that requires a clear and rigorous definition. Secondly, by means of formalization we may be able to unearth implicit assumptions which may not be entirely clear in the informal expositions, especially in the case of philosophical texts that are distant in time. Thirdly, by trying to translate the grammatical structure of ordinary language into the syntactic form of a logical one, new interesting questions about the role and use of philosophical terms may arise. There is also a last important reason — especially for the logicians — why we should formalize arguments: sometimes the use of technical resources does not help that much in understanding the arguments and assessing their validity, but it is crucial to improve our conception of logic itself. Sometimes classical logic does not provide enough technical background to justify the validity of arguments that we are inclined to accept for deep philosophical reasons. In this case, the exercise of formalization turns out to be an attempt to explore the possibility of reforming logic or to show the applicability of some non-classical formal inferences already at hand. Formalizing philosophical arguments may be an interesting way of going non-classical or shaping our general view about the role and the epistemic status of logical laws.

So much for the justification of the efforts presented in this paper. In section 2 I will expose Descartes' causal argument in an informal fashion, highlighting the most controversial assumptions and make a preliminary adaptation to the formalization. In section 3 I will discuss the logical axioms needed to implement my strategy of formalization. In section 4 I will present some metaphysical axioms needed to run the argument. In section 5 I will briefly examine the consequences of axioms for the

formal semantics, which will be a possible world semantics. In section 6 I will present the proof in the formal system which will be discussed from a philosophical point of view in section 7; section 8 is devoted to some conclusive remarks.

2. Informal exposition of the argument

There is a considerable debate on whether the Third Meditation contains two sorts of causal arguments or just two versions of one and the same argument, where the second version should be seen as a more explicative and detailed version of the first (Nolan 2014). It seems that Descartes himself suggested the former view in a letter to Mesland.[1] Since this paper is not actually concerned with exegetical issues, I will consider just the second formulation of the argument (or the second argument). Descartes' proof roughly runs along the following lines:

P1) I have in my mind the idea of God as a perfect infinite substance;
P2) Every idea in my mind must ultimately be caused by something having "actual existence" and not just existing in representation;
P3) The effect cannot be ontologically superior to the cause;
C) There exists a perfect and infinite substance which caused my idea of God.

The conclusion should follow from the fact that if I made up my idea of God, i.e. if this idea is caused by a finite substance as I am, then the effect — my idea of God — would be ontologically superior to the its cause, for a perfect and infinite substance — even if it has just "mental existence" — is something superior to a finite substance. In other words, there must be something real which provides enough representational content to my idea of God, for otherwise, the attributes which I lack in actual reality and I ascribe to God in my representation should have come out of nothing. Such exposition of the proof is quite vague and unclear, especially in regards to some metaphysical principles concerning the relations between reality and representation, an unspecified notion of ontological superiority, and causation. Descartes appeals to a significant deal of scholastic metaphysics to feel those gaps and make his argument clearer and more specific.

The first notion he briefly introduces is the distinction between formal and objective reality. The terminology is a bit puzzling, for by 'objective reality' he means the existence of something according to representation (which we would perhaps consider as a subjective reality), whereas by 'formal reality' he means "external existence", i.e. as a mind independent entity. For instance, I may say that Santa Claus has objective reality — given that I have a certain idea of Santa Claus as having certain properties — but no formal reality, since he does not exist over and above my representation. Therefore, the first premise of the argument (P1) may be rephrased

by saying that God has objective reality, i.e. he exists as an idea in my mind and that it is true according to my representation (i.e. true in objective reality) that he is an infinite substance.

The second issue is related to an alleged ontological difference between distinct kinds of entities. Descartes relies on the hierarchical picture of reality of medieval scholasticism, something known as the idea of the *Great Chain of Being* (Lovejoy 1936). In simple terms, not all entities exist in the same manner and with the same ontological dignity. Firstly, Descartes adopts the classical ontological categories of substance, accident, and mode.[2] Substances are more fundamental than modes and accidents, since modes and accidents are nothing but ways of being of substances. Among substances, infinite substances are somehow superior to finite ones, for they display some sort of perfection that finite substances lack. Descartes sometimes uses the expressions "to have more reality than" to denote such a priority in the hierarchy of reality. It is crucial to Descartes' argument that those ontological distinctions and categorizations apply also to what exists in objective reality, i.e. to my ideas:

> For beyond doubt those ideas that represent substances to me are something greater, and contain, if I may use the term, more 'objective reality' in themselves, than those that represent merely modes or accidents. And by the same token, the idea by which I conceive a supreme God, eternal, infinite, omniscient, all-powerful, and the creator of all things that exist beside himself, certainly has more objective reality in itself than those by which finite substances are represented. (Descartes 1984, p.29)

For instance, my idea of Santa Claus is the idea of finite substance and possesses "more objective reality" than my idea of Santa's property of 'having a white beard'. By the same token, my idea of God possesses more objective reality than my idea of Santa Claus, for the former is represented as an infinite substance while the latter is the representation of a finite one.

The third important clarification concerns the relation of causation between real entities (entities having formal reality) and ideas (entities having objective reality). Ideas always have an ultimate cause which must have formal reality. Even if certain ideas may be caused by other ideas, the causal chain must always end up to a mind independent cause. This is the only existential assumption of Descartes' argument, i.e. the only assumption that entails the existence of something having formal reality. Clearly, Descartes feels this assumption to be legitimate by the fact that he already knows that there must be at least one entity in formal reality: himself, as a finite thinking substance. The second important point regarding causation is how the causal mechanism works between formal and objective reality. Descartes conceives causation as a sort of fundamental property transmission according to which the effect cannot exceed its cause, otherwise there would be a creation "ex nihilo" of the effect's fundamental properties. Such an understanding also applies to causation

when conceived as a relation between things and ideas. As a consequence, a mode having formal reality cannot cause the idea of substance, a finite substance having formal reality cannot cause the idea of an infinite substance. Reality — formal or objective — cannot be "exceeded" through a causal relation.

Given these metaphysical remarks, we are in position to recast the causal argument:

A1) It is true according to my representation that God is an infinite substance;

A2) For every idea x in my representation there must always be something in reality which causes x;

A3) Everything that exists (in reality or representation) exists either as a finite substance or as an infinite substance, or as a mode;

A4) If x is a finite substance and x causes the idea of y, then it is true according to my representation that y is a finite substance or a mode;

A5) If x is a mode and x causes the idea of y, then it is true according to my representation that y is a mode;

A6) If x is a finite substance and x causes the idea of y, then it is true according to my representation that y exists;

C1) Neither a finite substance nor a mode in reality may be the cause of my idea of God;

C2) There must be — in reality — an actual infinite substance which causes my idea of God.

Where I have dropped the terminology objective/formal in favor of the more intuitive distinction between "true in reality" vs "true according to representation". This is the exposition of the causal argument to which I will refer in the following sections where the formalization will be developed.

3. The logical background

In this section I will start the formalization of Descartes' causal argument considering just the main logical issues related with this task. The formal rendering of the metaphysical assumptions will be dealt with in the next section.

I will start examining the issue of the propositional logic underlying the argument. According to the formulation provided in the previous section, it is immediately evident that we cannot run Descartes' proof in classical logic: expressions such as 'true in representation' or 'true in reality' indicate some sort of modality which is required to translate the premises into symbolic language. The adopted strategy is to consider a modal operator '□' to which we preliminary attribute the meaning of 'true accord-

ing to representation'. Clearly, the sort of modality represented by the box operator is not alethic (more on this later).

The weakest system of propositional modal logic is **K**, characterized by the axiom of distribution:

(DIST) $\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$

and the rule (NEC) of necessitation:

$$\frac{\vdash \alpha}{\vdash \Box\alpha}$$

When the box is interpreted as a knowledge operator (e.g. in epistemic logic), distribution and necessitation entail "logical omniscience", i.e. the fact that if an epistemic agent S knows that p, then S knows all logical consequences of p. In our case we may speak of a sort of "logical omni-representability", according to which if it is true in my representation that p, then all logical consequences of p are also true according to my representation. As it will be shown, this fact is required if Descartes' argument is to be valid; in section 7 I will discuss the adequacy of logical omni-representability to the general epistemological setting of the Third Meditation.

The system **K** is too weak to implement Descartes' proof, thus possible extensions need to be discussed. It is easy to see that the axiom (M):

(M) $\Box\alpha \rightarrow \alpha$

while acceptable in aletheic modal logic, is untenable when we ascribe to the modal operator the meaning of 'true in representation'. Indeed, the whole point of Descartes' argument is to show that God exists while maintaining the hypothesis of the deceiving demon; and (M) precisely says that my representation never deceives me (what is true according to my representation must be true in reality). The weaker "deontic axiom" (D) seems more palatable:

(D) $\Box\alpha \rightarrow \neg\Box\neg\alpha$

which says that if α is true according to my representation, then $\neg\alpha$ cannot be also true in my representation. In other words, (D) says that I cannot have contradictory representations, i.e. my ideas form a whole consistent body of "mental images". It is worth saying that (D) is indispensable to construct Descartes' proof; however, one may object that if I am subjected to the power of a deceiving demon, then I may develop an inconsistent image of reality in my own mind. I would preliminarily say that it is not clear in what sense I may have inconsistent representations: representation has to do with conceivability, and it seems counterintuitive to suppose that I am able to conceive contradictory states of affairs.

The weakest propositional system in which we may mount Descartes' proof is the system **D**, obtained by extending **K** with the axiom (D). However, one may consider the possibility of further extensions of **D** in order to capture a broader philosophical background compatible with the train of thought of Descartes' *Meditations*. For instance, it would be rational to accept the modal axioms (4) and (5):

(4) $\Box\alpha \to \Box\Box\alpha$

(5) $\neg\Box\neg\alpha \to \Box\neg\Box\neg\alpha$

Axiom (4) may be read as a sort of reflexivity of representation: if α is true according to my representation, then I represent the state of affairs that α is true according to my representation as true. The same sort of reflexivity applies to the case in which I do not represent α as false. If we are to endorse these principles as capturing the meaning of the "representation of operator" (i.e. the box), then the system we shall adopt is $\mathbf{D}_5$ (i.e. **D** + (4) + (5)).

The next step is to evaluate the quantificational logic underlying the argument. We may start by adding the classical first-order axioms to the system **D**:

(Q1) $\forall x(\alpha \to \beta) \to (\forall x\alpha \to \forall x\beta)$

(Q2) $\forall x\alpha \to \alpha[t/x]$ for every term 't'

and the rule (GEN) of generalization:

$$\frac{\alpha}{\forall x\alpha}$$

Call **CQD** (classical quantificational **D**) the system obtained by adding (Q1), (Q2), and the rule of generalization to **D**. It is easy to see that **CQD** is too strong, for it is possible in it to prove the Converse Barcan Formula:

(CBF) $\Box\forall x\alpha \to \forall x\Box\alpha$

(CBF) says that — for instance — if it is true in representation that everything ages, then for every x existing in formal or external reality, it is true in representation that x ages. The validity of (CBF) entails that the domain of quantification of what exists in representation must include the domain of quantification of external reality. In other words, this would imply that I have an idea of everything that exists independently of my mind. It is clearly unacceptable — from the perspective of Descartes' *Meditations* — to endorse such a wide quantificational power of my mind; hence, we need a quantificational version of **D** weaker than **QCD**.

The only way to construct a system of quantificational modal logic in which the converse Barcan formula is not a theorem is to use the axioms of free logic (Gabbay

and Guenthner 2010). As a matter of fact, (CBF) follows from the axiom of distribution, the axiom (Q2) and the rule of generalization; yet (Q2) is not generally valid in free logic, so the derivation of (CBF) is blocked. This fact ensures the independence of the domain of quantification in representation from the domain of quantification in actual reality, and thus the possibility of being deceived about the ontological inventory of reality.

In a nutshell, the basic idea of free logic is to reject the implicit assumption of classical quantificational logic according to which every closed term must have a denotation which belongs to the domain of quantification. This assumption has been considered too strong for it entails an ontological commitment of logic: the domain of quantification cannot be empty and all closed terms must have an existent denotation. On the contrary, in free logic the fact that our language contains closed terms does not entail existential statements regarding their denotations and there may be models of the axioms with empty domain.

In the context of providing a logical background to Descartes' train of thought in the Third Meditation, it is not easy to assess whether free logic is an adequate selection of a logical theory. Yet I may attempt to show at least the compatibility of this selection with Descartes' epistemic background. While meditating on what may exist outside of my mind and contemplating the possibility of being deceived by the demon, I find a variety of ideas and ways of representing things as contents of my thought. For each idea (except the idea that I have of myself as a thinking substance) I cannot possibly know whether they represent something existing in the external world. And yet I am able to use singular terms and predicates to make assertions. This means that I am innocent of whether the terms I am using have a denotation which belong to the domain of quantification of "formal reality" (i.e. external world) or just to the domain of quantification of "objective reality" (i.e. my representation). Thus I may interpret the use of free logic as standing for the following epistemic situation: some terms may lack reference in the external world, and yet I may use them to make true or false assertions, for they may refer to an idea in my mind. As a consequence, in being allowed to use allegedly non referring terms (at least lacking a reference in formal reality) to make assertions in thought, I am implicitly adopting a certain sort of free logic.

The adoption of free logic has also an interesting philosophical consequence: it makes explicit the relations between the causal argument of the Third Meditation and the "cogito". To understand this fact we must bear in mind that in free logic, in general, we cannot instantiate an universal sentence for a singular term 't'. This is precisely due to the fact that 't' may be an empty term (or a term lacking reference in the external world) and thus it cannot be interpreted as an element of the domain of quantification. Therefore, before instantiating a universal sentence using a term 't', we must show that 't' has a reference in formal reality (or a reference existing

in the external world). For instance, if we assert that everything is a substance or a mode, we need to show that there is at least one entity in the external world to grant that our assertion is not vacuously true (empty domain). However, given that I may be possibly deceived on everything by the demon, what guarantees do I have that certain general principles (such as 'everything is a substance or a mode') are not uninteresting platitudes of an empty universe? Here the "cogito" comes into play: for given that I cannot doubt my existence as a mind, I do have access to an existential truth (true according to the external world). It is only by virtue of the fact that I "see" the truth of the sentence 'I exist' that any universal statement is informative about the external world (which still may include only one entity). In other words the "cogito" fills an existential gap which — without the use of free logic — cannot be explicit, for it would be hidden in the metalinguistic assumptions of the logical system. Free logic makes explicit the role and the significance of the "cogito".

From a technical point of view, there is the issue of selecting a particular axiomatization, for there are three main systems of free logic: positive, negative, and neutral. Positive free logic allows the possibility of sentences containing empty terms to be true or false; according to negative free logic they are always false; neutral free logic allows for truth value gaps. If we are to interpret sentences containing empty terms (i.e. terms not having a reference in the external world) as sentences talking about the domain of representations, then it makes sense to adopt positive free logic, for 'Santa Claus is bearded' is true according to my representation in spite of the fact that Santa Claus does not exist in the external world. As we will see, such a choice is not uncontroversial.

I will focus on the minimal axiomatization needed to run Descartes' causal argument. Given that I will not make use of an identity predicate, I will consider the axiomatization of positive free logic without the axioms of identity. As a consequence, it will not be possible to use the existential quantifier to construct formulas that says that 't exists' for a term 't' (e.g. '$\exists x(x = t)$'). In this case, the language of free logic must include an existence predicate 'E' as a primitive symbol. The axioms (and schemas) for the first-order fragment of the adopted logical system are the following:

(Q1) $\forall x(\alpha \rightarrow \beta) \rightarrow (\forall x \alpha \rightarrow \forall x \beta)$

(FQ2) $\forall x \alpha \rightarrow (Et \rightarrow (\alpha[t/x]))$ for every term 't'

(FQ3) $\forall x Ex$

which makes clear in which respect our system differs from its classical counterpart: instantiation of universals requires an additional existential premise (FQ2), and the axiom saying that the quantifier ranges over the domain of existing entities (FQ3). I call **FQD** the system obtained by adding (Q1), (FQ2), (FQ3), and the rule of generalization to **D**. **FQD** is thus a system of modal first-order positive free logic: the weakest

set of axioms and rules that allows us to run Descartes' causal argument in a formal fashion.

A last remark on language: we have considered axioms for a propositional signature including only '¬' and '→', and a first-order signature including only '∀'. In the following the additional symbols '∨', '∧', '∃' will be used as customary abbreviations.

With our logical theory in place, it is time to translate the main metaphysical assumptions of Descartes' argument into the symbolic language. This will be the scope of the next section.

4. The metaphysical background

As made clear in section 2, Descartes' argument relies on a series of assumptions concerning the metaphysical background of the Third Meditation; these assumptions are concerned with three main metaphysical notions: 1) the hierarchical structure of Being, 2) the idea of God in my mind, and 3) the relation of causation, especially between things and ideas. In this section I will present three groups of axioms corresponding to each one of these issues to be added to the logical system **FQD** as the metaphysical background theory.

I will start by considering the ontological axioms which introduce the main categories of scholastic metaphysics. There are three main ontological categories: infinite substance, substance, and mode. We need to formalize two main ideas:

1. That everything which exists (in external reality) is either an infinite substance, a finite substance, or a mode;
2. That these categories are mutually exclusive, i.e. if something is a mode cannot be a substance, etc. . .

Considering that our formal language must have an existence predicate (in absence of identity), a nice way to formalize 1 and 2 would be to introduce specific existence predicates: the symbol 'E^{IF}', standing for 'existing as a infinite substance', 'E^{FS}' which may be read as 'existing as finite substance', and 'E^M' which shall be interpreted as 'existing as a mode'. We will consider these three predicates as primitive symbols of our language and introduce the existential predicate 'E' by means of the following abbreviation:

(EX)   'Et' abbreviates '$(E^{IS}t \lor (E^{FS}t \lor E^M t))$' for every term '$t$'

In other words, to say that something exists just amounts to saying that something exists as an infinite substance, as a finite one, or as a mode. Clearly, the predicate 'E' in axioms (FQ1) and (FQ2) shall be considered as abbreviating the expression in

(EX). The next step is to introduce the idea of mutual exclusion between ontological categories; this can be easily done by the following three axioms:

(E-IS) $E^{IS}t \rightarrow (\neg E^{FS}t \wedge \neg E^{M}t)$, for every term '$t$'

(E-FS) $E^{FS}t \rightarrow (\neg E^{IS}t \wedge \neg E^{M}t)$, for every term '$t$'

(E-M) $E^{M}t \rightarrow (\neg E^{IS}t \wedge \neg E^{FS}t)$, for every term '$t$'

Notice that the axioms are not formulated as universally quantified sentences, yet they are schemas valid for every term. This is because they are vacuously true in the case of a term 't' lacking reference in formal reality. Indeed, if 't' refers to a nonexistent object (i.e. something existing only in representation), then the antecedents in (E-IS), (E-FS), and (E-M) are false, thus the conditionals are true. The main consequence of this fact is that these axioms may be instantiated using any term without requiring the additional condition of the existence in formal reality of the referent of the term at issue (as in the case of universally quantified sentences).

Successively, we introduce an axiom saying that I have in my mind a representation of God as an infinite substance; this corresponds to assumption (A1) in the informal exposition of the argument. The axiom of the idea of God is the following:

(IG) $\square E^{IS}g$

where the singular term 'g' is the symbolic representation of the proper name 'God'.

The last issue concerns the axiomatization of the relation of causation between things and ideas, symbolically represented by a two place predicate 'C'; hence, the expression 'Cxy' shall be interpreted as 'x causes the idea of y'. We need axioms putting restrictions on what sort of idea may be caused respectively by an infinite substance, a finite one, and a mode. More specifically, we will introduce three axioms formalizing assumptions (A4), (A5), and (A6) in the informal exposition of the argument:

(C-IS) $E^{IS}k \rightarrow (Ckt \rightarrow (\square Et))$, for every terms '$k$' and '$t$';

(C-FS) $E^{FS}k \rightarrow (Ckt \rightarrow (\square E^{FS}t \vee \square E^{M}t))$, for every terms '$k$' and '$t$'

(C-M) $E^{M}k \rightarrow (Ckt \rightarrow (\square E^{M}t))$, for every terms '$k$' and '$t$'

Axiom (C-IS) says that if k is an infinite substance and k causes the idea of t, then t exists in representation (either as an infinite substance, as a finite one, or as a mode). Axiom (C-FS) says that a finite substance can cause only ideas of finite substances or modes; axiom (C-M) says that modes can only cause ideas of modes. Notice that also these axioms are schemas; given that they have an existential antecedent, they are true (vacuously) even in the case of terms lacking a reference in formal reality.

To complete the axiomatization of the relation of causation between things and ideas, the formalization of assumption (A2) is needed, i.e. an axiom saying that every idea must have an ultimate cause having "formal existence", namely being an actual entity independent of my mind. This is the only existential axiom that will be explicitly used in the proof of the existence of God (section 6):

(C-EX) $\exists x C x t$, for every term 't'

Albeit not necessary to prove the existence of God, the axiomatic system of Descartes' meditations may be enriched with another existential axiom which may be called the 'cogito axiom', saying that I exist as a finite (thinking) substance:

(COGITO) $E^{FS}m$

Where the singular term 'm' stands for myself (or the meditator). We call **MED** the formal system of Descartes' meditations, obtained by adding to **FQD** the axioms of the existential predicates ((E-IS), (E-FS), and (E-M)), the axiom of the idea of God (IG), the axioms of causality between things and ideas ((C-IS), (C-FS), (C-M), and (C-EX)) and the cogito axiom (COGITO). In section 6 I will present the proof of the fact that God has formal reality (i.e. exists in the external world) in the system **MED**. In section 5 I will develop some considerations on the formal semantics for **MED** and its models.

5. Semantics

In this section I will present the formal semantics for the logical theory underlying the system MED (and some possible extensions) and briefly discuss some interesting aspects of its models. In order to manage the complexity of the system, models for the system of Meditations will be built in two steps: 1) models of the propositional fragment, 2) models of its first-order extensions.

As clarified in section 3, the propositional logic underlying **MED** is **D** (envisaging the possibility of extending it to $\mathbf{D_5}$). Therefore, Kripke's possible worlds semantics will be implied. Before disputing the details of the effective conditions on the accessibility relation, it is worth presenting the intended model for the Meditations. Intuitively, we may admit two "possible" worlds: one corresponding to 'formal reality' — i.e. the external actual world — and the other to 'objective reality' — i.e. the world of my representation. These worlds will be respectively labeled as $\mathbf{w}_f$ and $\mathbf{w}_o$. The main desiderata on the accessibility relation are:

- The external world $\mathbf{w}_f$ must not access itself, for we want to contemplate the possibility for my representation to be wrong about the actual world. Thus, it

should be possible that for some sentence p, $\Box p$ is true according to $\mathbf{w}_f$ (i.e. p is true in representation) while p is false according to $\mathbf{w}_f$ (i.e. p is false in the actual world). As previously said, this possibility is allowed by rejecting the axiom (M);

- The external world $\mathbf{w}_f$ should not be accessed by any other world, for it is desirable — by virtue of the hypothesis of the deceiving demon — to maintain truth in representation (i.e. truth of sentences starting with the box operator) independent from truth in the external world. If the external world $\mathbf{w}_f$ were accessible from a distinct world w, then the truth of a sentence p in $\mathbf{w}_f$ would imply the truth of $\Box p$ in w (i.e. truth in representation according to w); in other words, the obtaining of a certain state of affairs in the external world would imply its true representation in a certain world;

- The representational world $\mathbf{w}_o$ must be accessible from the external world $\mathbf{w}_f$, for if it is actually true that $\Box p$ (i.e. that it is true in the external world that p is true according to my representation), then p must be true in the representational world. Otherwise, it may be the case that $\Box p$ is actually true while p is false in the representational world $\mathbf{w}_o$;

- The representational world $\mathbf{w}_o$ must access itself, for we want assertion of truth in representation to entail truth in the representational world. Thus if $\Box p$ is true in $\mathbf{w}_o$, this shall entail that p is true in $\mathbf{w}_o$.

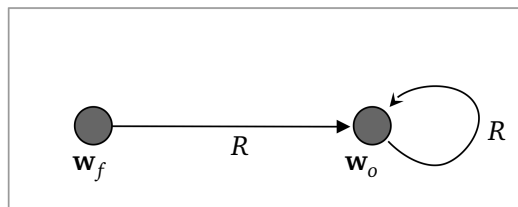We may graphically represent our basic intended model of the propositional fragment of our system:

Figure 1: Intended model of Meditations

where 'R' denotes the accessibility relation.

Clearly, the requirement of a set of axioms which uniquely picks out our intended model will be strong. Thus our first question will be: is the intended model a model of **D**? To answer the condition that the axiom (D) imposes on the accessibility relation must be considered; let **W** be the set of all possible worlds; the axiom (D) corresponds to the following semantic clause:

(D*) for every w in **W**, there is u in **W** such that wRu

In other words, every world must access some world. This condition is clearly sat-
isfied by the intended model, for $\mathbf{w}_f$ access $\mathbf{w}_o$ and $\mathbf{w}_o$ access itself. However, it is
worth considering also unintended models, in order to understand to what extent the
system D captures our intuitive idea of the epistemic background of the Meditations.
The most unwelcome unintended models seem to be "branching models" where the
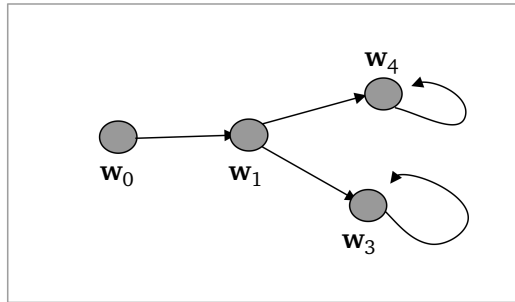accessibility relation may be graphically represented in the following way:

Figure 2: "Branching" model of D.

According to the previous considerations, world $\mathbf{w}_0$ must be the external world,
worlds $\mathbf{w}_2$ and $\mathbf{w}_3$ representational worlds, the classification of world $\mathbf{w}_1$ is a bit more
problematic. Moreover, the model of Figure 2 is somehow "unorthodox" insofar as if
p is true in representation according to the external world $\mathbf{w}_0$, the representational
worlds may not agree with $\mathbf{w}_0$ on the truth in representation of p. Indeed, it is com-
patible with the considered model that p is true in $\mathbf{w}_0$ and $\mathbf{w}_1$ (which makes $\Box p$ true
in $\mathbf{w}_0$) while false in $\mathbf{w}_2$ and true $\mathbf{w}_3$ (which makes $\Box p$ true in $\mathbf{w}_3$ and false in $\mathbf{w}_2$).
If this is to be considered a fundamental requirement, then branching models must
be avoided.

Branching models may be avoided by adding axioms (4) and (5). By adopting the
system $\mathbf{D}_5$ we will have more intuitive models, i.e. models where when there is one
world non auto-accessible, it is easier to interpret the remaining worlds as possible
ways of representing reality, for they all agree with the external world on what is true
in representation. To see how this fact holds, consider the requirement imposed on
the accessibility relation by (4) and (5):

(4*) for every $w, v.u$ in $\mathbf{W}$, if wRu and uRv, then wRv

(5*) for every $w, v.u$ in $\mathbf{W}$, if wRu and wRv, then uRv

By applying clauses (4*) and (5*), the branching model of Figure 2 turns out to be
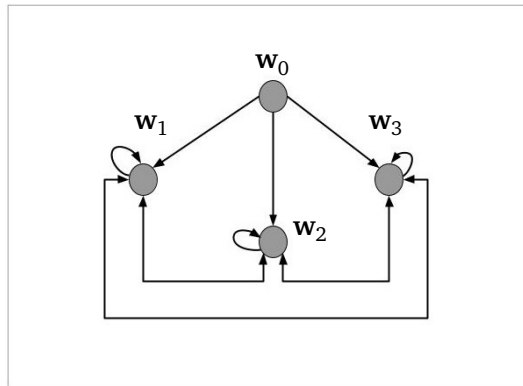as follows:

Figure 3: Avoiding branching in $\mathbf{D}_5$

where it is easy to see that $\mathbf{w}_0$ should count as the external world and the other worlds as representational. In general, in the case of models with more than two worlds (as in Figure 3), the accessed worlds may be interpreted as ways of representing the external world. Those worlds agree on truths in representation, i.e. on the truth values of sentences starting with the box operator. Every sentence of the form $\Box p$ may be interpreted as saying that p is true under all possible ways of representing the external world; perhaps, we may consider truths of this sort as correspondent to Descartes' clear and distinct ideas, i.e representations which are the same independently on the way we access to them in our minds (truths in all ways of representations). Under this interpretation, the existence of God must ensure the truth of the axiom (M) which may be taken to say that every clear and distinct representation is adequate to the external world (i.e. every sentence p true in every way of representation is true in the external world, '$\Box p \rightarrow p$').

The main conclusion is that by adopting $\mathbf{D}_5$ instead of $\mathbf{D}$ there is a more natural interpretation even of models with more than two worlds, provided that there is at least one world not accessed by any other.

So much for the discussion of the propositional fragment. Regarding the extension of $\mathbf{D}$ (or $\mathbf{D}_5$) to first-order, we have adopted positive free logic; therefore, some remarks on formal semantics are needed. In general, a structure M for $\mathbf{FQD}$ has the following components:

$$M = \langle \mathbf{W}, R, D, dom, v \rangle$$

Where

1. $\mathbf{W}$ is the set of all possible worlds;
2. R is the accessibility relation (which must satisfy the previously mentioned requirements) ;

3. D is the general domain over which all individual constants are rigidly interpreted;

4. *dom* is the domain function, i.e. the function that associate to every w in **W**, the domain $dom(w) \subseteq D$;

5. ν is a valuation function which receives an element of **W** and a closed formula as arguments and returns a truth value in $\{0, 1\}$;

6. For every w in **W**, we also have an interpretation function $(\dots)^w$ which picks out the extensions of predicates and terms.

Since we are using free logic, for some possible world w in **W**, $dom(w)$ may possibly be the empty set. Moreover, positive semantics has been chosen, thus for every possible world w in **W**, two domains are needed: the internal domain — $dom(w)$ — over which first-order quantifiers range, and the external domain — $D/dom(w)$ — in which empty terms are interpreted. As shown, another consequence of adopting free logic is that neither the Barcan formula nor its converse hold, thus given two worlds w, u in **W** such that uRw, there is no constraint on $dom(u)$ and $dom(w)$, i.e. neither $dom(u)$ must be a subset of $dom(w)$ nor $dom(w)$ must be a subset of $dom(w)$.

Having described the semantic setting and the axiomatic theory it is now time to formalize Descartes' causal argument. This will be the topic of the next section.

6. The formal proof

In this section the crucial part of this paper is developed: the formal proof of the existence of God in the system **MED**. Firstly, it must be clarified what will be proved: *the existence of an infinite substance which causes my idea of God*. In symbolic terms:

(G) $\exists x (Cxg \wedge E^{IS} x)$

To show how (G) follows from the axioms of **MED**, two important theorems must be proved:

Theorem 1. *No finite substance causes my idea of God, i.e.*

$$\forall x (Cxg \to \neg E^{FS} x)$$

Theorem 2. *No mode causes my idea of God, i.e.*

$$\forall x (Cxg \to \neg E^{M} x)$$

Proof of **Theorem 1**:

1. $E^{FS}x \rightarrow (Cxg \rightarrow (\Box E^{FS}g \vee \Box E^{M}g))$ (by schema (C-FS) for the term 'x')
2. $E^{FS}x$ (assumption)
3. $Cxg \rightarrow (\Box E^{FS}g \vee \Box E^{M}g)$ (by Modus Ponens on 1,2)
4. $\neg(\Box E^{FS}g \vee \Box E^{M}g) \rightarrow \neg Cxg$ (by contraposition on 3)
5. $(\neg\Box E^{FS}g \wedge \neg\Box E^{M}g) \rightarrow \neg Cxg$ (by De Morgan's theorem)
6. $E^{IS}g \rightarrow (\neg E^{FS}g \wedge \neg E^{M}g)$ (by schema (E-IS) for the term 'x')
7. $\Box E^{IS}g \rightarrow \Box(\neg E^{FS}g \wedge \neg E^{M}g)$ (by (NEC) and (DIST)) on 6)
8. $\Box E^{IS}g$ (Axiom (IG))
9. $\Box(\neg E^{FS}g \wedge \neg E^{M}g)$ (by Modus Ponens on 7,8)
10. $(\Box\neg E^{FS}g \wedge \Box\neg E^{M}g)$ (propositional tautology of **D**)
11. $\Box\neg E^{FS}g$ (eliminating the conjunction in 10)
12. $\Box\neg E^{FS}g \rightarrow \neg\Box E^{FS}g$ (instance of axom (D))
13. $\neg\Box E^{FS}g$ (by Modus Ponens on 11, 12)
14. $\Box\neg E^{M}g$ (eliminating the conjunction in 10)
15. $\Box\neg E^{M}g \rightarrow \neg\Box E^{M}g$ (instance of axiom (D))
16. $\neg\Box E^{M}g$ (by Modus Ponens on 14,15)
17. $(\neg\Box E^{FS}g \wedge \neg\Box E^{M}g)$ (by introducing the conj. on 13,16)
18. $\neg Cxg$ (by Modus Ponens on 5,17)
19. $E^{FS}x \rightarrow \neg Cxg$ (discharging assumption 2)
20. $\forall x(E^{FS}x \rightarrow \neg Cxg)$ (by GEN)

The proof of theorem 2 runs along the same lines and will be omitted. From Theorem 1 using the "cogito axiom" it is easy to prove a corollary saying that I cannot be the cause of my idea of God, i.e. that the idea of God cannot be a fictitious one. Even though this corollary is an intermediate step in Descartes' strategy, it may be bypassed (as in the informal exposition of the argument of section 2).

Before proving (G), it is worth mentioning that the following sequents are valid in free logic:

(*) $\forall x\alpha, \exists x\beta \vdash \exists x(\alpha \wedge \beta)$

(**) $\forall x(\alpha \rightarrow \beta), \forall x(\alpha \rightarrow \gamma) \vdash \forall x(\alpha \rightarrow (\beta \wedge \gamma))$

(***) $\exists x(\alpha \wedge (\alpha \rightarrow \beta)) \vdash \exists x(\alpha \wedge \beta)$

It is now possible to prove

Theorem 3. *Exists an Infinite Substance which is the cause of my idea of God, i.e. (G).*

Proof of (G):

1. $\forall x(Cxg \rightarrow \neg E^{FS}x)$ (by Theorem 1)
2. $\forall x(Cxg \rightarrow \neg E^{M}x)$ (by Theorem 2)
3. $\forall x(Cxg \rightarrow (\neg E^{FS}x \wedge \neg E^{M}x))$ (by (**) on 1 and 2)
4. $\exists x Cxg$ (by (C-EX) instantiated for 'g')

5.	$\exists x(Cxg \wedge (Cxg \rightarrow (\neg E^{FS}x \wedge \neg E^{M}x)))$	(by (*) on 3,4)
6.	$\exists x(Cxg \wedge (\neg E^{FS}x \wedge \neg E^{M}x))$	(by (***) on 5)
7.	$\forall x Ex$	(axiom (FQ3))
8.	$\exists x(Cxg \wedge (Ex \wedge (\neg E^{FS}x \wedge \neg E^{M}x)))$	(by (*) on 6,7)
9.	$E^{IS}x \leftrightarrow (Ex \wedge (\neg E^{FS}x \wedge \neg E^{M}x)$	(by definition of 'E')
10.	$\exists x(Cxg \wedge E^{IS}x)$	(by 9 and substitution of equivalents in 8)

Verdict: the argument is valid in **MED** which is possibly the weakest system of quantified modal logic in which the presented proof may be mounted. Conceptual issues related to the philosophical import of **MED** and the soundness of the argument will be discussed in the next section.

7. Philosophical remarks on the formalization

In this section I will discuss some philosophical issues arising both from the presented formalization of Descartes' causal argument and from the structure of the argument itself. Given the validity of the argument, the purpose of the present discussion is to assess its soundness and the import on the argument of the philosophical assumptions underlying the general setting of modal free logic. In particular, three issues will be addressed:

1. the problem of "metaphysical omniscience";
2. the understanding of rigid designation;
3. the behaviour of empty names.

Metaphysical omniscience. This issue has to do with the problem of logical omniscience — briefly mentioned in section 3. The simplest way to show how this problem is engendered is to consider step 7 in the proof of Theorem 1, in which the following inference is implicitly made:

$$\frac{\vdash E^{IS}g \rightarrow (\neg E^{FS}g \wedge \neg E^{M}g)}{\Box(E^{IS}g \rightarrow (\neg E^{FS}g \wedge \neg E^{M}g))}$$

The inference is justified by both the rule of necessitation and the fact that the formula '$E^{IS}g \rightarrow (\neg E^{FS}g \wedge \neg E^{M}g)$' is a valid formula of **MED**; and it is indeed, for it is an instance of the axiom (E-IS). The use of the rule of necessitation entails that the meditating subject must recognize the truth in representation of all the axioms of the system (or premises of the argument), i.e. of the metaphysical principles used in the proof. As a consequence, the fact that **MED** is based on the system **K** results in a

metaphysical omniscience of the meditating subject who, in spite of acknowledging the possibility of being deceived on the evidence that '$2 + 2 = 4$', cannot be deceived on the fundamental principles of scholastic metaphysics and — by the combination of the rule of necessitation with the axiom of distribution — on all of their logical consequences. This is a quite controversial matter; for even if we may have reasons to argue in favor of logical omniscience and of the fact that the demon cannot deceive me on the basic principles of logic, it is hardly imaginable that similar arguments may be formulated to justify a metaphysical omniscience. After all, why is scholastic metaphysics immune to the evil manipulation and control of the deceiving demon? Why should it be more evident than the simplest truth of mathematics such as '$2+2 = 4$' or 'A triangle has three angles'?

This is no novelty in the history of objections to Descartes' meditations. What is interesting here is how the adopted strategy of formalization makes the difficulty clear and evident: by connecting the problem of the status of principles of first philosophy to the problem of propagation of the box operator through the consequence relation. In other words, a well known and discussed problem of epistemic modality — i.e. the problem of logical omniscience — turns out to reappear in Descartes' argument as the problem of metaphysical omniscience casting reasonable doubts on both the status and the use of scholastic metaphysics in Descartes' epistemological project. After all, the whole point of this project was to determine the principles of "First Philosophy" by applying the methodic doubt and thus by accepting the possibility that even the most fundamental metaphysical truths may not be recognized as such in the representation of the meditator. Clearly, the "cogito" survives the induced skepticism and my existence is true both in external reality and representation. However, as shown by the formalization, the "cogito axiom" plays no explicit role; in this respect, Descartes' argument is not different from the classical ontological arguments in the Scholastic tradition.

Verdict: *the argument is valid, yet there are good reasons to think that it is not sound. This is because it requires the assumption that first metaphysical principles are not just true but also acknowledged as true in representation by the meditator, surprisingly surviving the methodic doubt without a clear explanation.*

The understanding of rigid designation. The system **MED** is based on a language that does not contain the description operator and on Kripke's semantics; therefore, all singular closed terms are rigid designators. Yet how rigid designation is to be understood in the framework of Descartes' meditations, a philosophical project which is completely innocent of the debate on referential behaviour of proper names across possible worlds? In this respect, the proposed formalization introduces new philosophical elements which may or may not be in harmony with the spirit and the context of the causal argument. I will start by discussing the issue for terms referring to

actually existing individuals; the issue of empty names will be discussed later on.

The basic idea of rigid designation combined with Kripke's conception of possible worlds (Kripke 1980) offers a natural interpretation of the behaviour of proper names in the context of Descartes' meditations. Recall that according to the proposed strategy, possible worlds are considered either as external reality or as representational domains. This seems to engender no tension with Kripke's account of merely possible worlds as representational devices, for the fact that the sentence 'Socrates is wise' is true in the possible world **W** amounts to the fact that "the actual Socrates" is represented as wise in the world **W**. In our setting, to say that 'Socrates is wise' is true in the objective reality $\mathbf{w}_o$, amounts to saying that "the actual Socrates" is represented as wise in my mind, i.e. in the representational device $\mathbf{w}_o$. As a consequence, statements true in representation of an individual t are statements about the actual individual t (i.e. existing in formal reality) and they are true in virtue of the way "the actual t" is represented in my mind. In conclusion, I may say that the implementation of Descartes' proof in modal logic carries no particular difficulty insofar as an *abstractionist conception* of possible worlds is adopted — i.e. possible worlds are not "distant planets" (Kripke 1980) yet just representational devices; and this seems to be the standard way of considering possible worlds when Kripke's semantics is used.

The behaviour of empty names. The previous analysis of truth conditions of sentences containing singular terms presents some difficulties when "empty names" are considered, i.e. singular terms which lack a reference in the external world. Given that free logic has been used in the formalization, the discussion of this topic is indispensable.

Consider the sentence 'Pegasus flies' and suppose it is true according to my representation; clearly we cannot spell out the truth conditions of 'Pegasus flies' as "it is true of the actual Pegasus that it is represented in my mind as flying". Formal semantics of positive free logic offers a nice mathematical tool to deal with this sort of situations, for an additional domain is considered where "fake references" for empty names are allowed; the main consequence is that we may formally treat empty names without substantial alterations of the main semantic clauses. Nevertheless, a philosophical interpretation — consistent with the context of Meditations — of the behaviour of empty names is desirable. The most natural way of interpreting empty names is to assume that they have no reference in the external world and that they rigidly refer to "ideas" (i.e. representations) of my mind. Hence, 'Pegasus flies' is true according to my representation iff the actual existing idea of Pegasus is "present" in my mind as the idea of a flying object. In a certain sense, all terms must have a reference and some terms are said to be empty only with respect to the external world: they refer to mere representations. Clearly, this account is highly controversial from our contemporary perspective, for it partially endorses a sort of *semantic idealism*

whose difficulties have been presented even before the famous "linguistic turn" (Mill 1843). However, any attempt to apply the conclusions of the huge debate on empty names in contemporary philosophy of language to the context of Descartes' meditation would take us too far; I will limit myself to expose some preliminary remarks on this topic.

It is reasonable to assume that Descartes would have endorsed a semantic idealism, at least insofar as names of fictional entities are concerned (as many modern philosophers did). With this premise in mind, our problem is if the proposed axiomatic system (and thus the argument itself) is consistent with the mentioned view on empty names. In more technical terms, it must be ensured that every singular term whose reference does not belong to the domain of quantification of the external world (i.e. formal reality) must have a reference belonging to the domain of quantification of the representational world (i.e. objective reality). This amounts to the validity of the following formula:

(EN) $\neg Et \rightarrow \Box Et$

for every term 't'. (EN) is not generally valid in **MED**, thus the possibility of using it as an additional axiom must be examined. To evaluate the acceptability of (EN) it is worth considering one of its consequences. Consider the following proof:

1. $\Box\neg Et \rightarrow \Box\Box Et$ (by (NEC) and (DIST) on (EN))
2. $\Box\Box Et \rightarrow \Box Et$ (propositional tautology of D5)
3. $\Box\neg Et \rightarrow \Box Et$ (by 1 and 2)
4. $\Box\neg Et \rightarrow \neg\Box Et$ (instance of (D))
5. $\neg\Box\neg Et$ (by the reductio ad absurdum on 2 and 3)

Clearly the proof cannot be carried out in **MED**, for it requires an underlying propositional logic stronger than **D**. Yet, in order to avoid some unintended models, we may have independent reasons to strengthen **MED** using $\mathbf{D_5}$. As a result, adding (EN) as an axiom entails that every negative existential is false in representation of every individual that can be named. This is plainly consistent with our previous discussion: if every proper name rigidly refers to either an actual entity or an idea, then from the perspective of my representation there are no empty names. And yet this result sounds extremely awkward: how could it be that there are no negative existentials true according to my representation? According to Descartes, when I utter in my mind that Pegasus does not exist, I am actually saying that Pegasus has no formal reality, even though the idea of Pegasus exists in my mind (i.e. Pegasus has objective reality). Yet the considered formal language has not enough expressive power to translate this sort of statements: for when the existential predicate is inside the scope of the box it automatically means 'existence in representation'. In other words, it is impossible to formalize sentences stating existence or nonexistence in

the external world from the perspective of representation. Considered such syntactic limitations, the adoption of both (EN) and semantic idealism for empty names may be too controversial.

There is an interesting lesson to be learned here: by using quantified modal logic to formalize Descartes' argument a big deal of philosophy of language has been implicitly imported in the general philosophical context of the discussion. As a result, theoretical tensions unsurprisingly arise, especially given the innocence of early modern philosophers of the main difficulties related to both negative existentials and empty names. I regard this fact as an additional reason why formalizing classical arguments of the philosophical tradition has its intrinsic value: for it unearths a big deal of philosophical problems which are not immediately evident in the informal exposition of those arguments.

8. Conclusion

I have presented an attempt to formalize Descartes' causal argument for the existence of God in quantified modal logic, using the axioms of free logic for the first-order fragment of the underlying theory. There are two last interesting remarks on the proposed formalization. The first remark has to do with the relations between metaphysics and logic. The reader should have noticed that a big deal of metaphysics has been imported in the logical theory, for instance, in the case of the axioms governing the behavior of the existential predicates. Ontological categories such as 'substance' and 'mode' are included even in the logical vocabulary, for corresponding existential predicates have been used. This may be considered as unacceptable from the perspective of the neutrality of logic. Besides the discussion of this issue — which will take us too far — it is interesting to note that by attempting such formal reconstructions of classical arguments contemporary problems of philosophical logic — such as the topic neutrality of logic — gain force and clarity, for their philosophical urge may be understood more profoundly.

The second remark has to do with the technical complexity of the formalization. It seems quite evident that the present paper is not accessible to the philosophy students with a basic logical background (e.g. propositional and first-order classical logic). I regard this fact as the main reason why study of "advanced and non-classical logics" should be fostered in philosophy courses: for the philosophical discourse with its dialectical complexity goes far beyond classical logic; and logic itself, becomes more and more "philosophical" as the complexity of the considered systems increases.

References

Descartes, R. 1984. *The Philosophical Writings of Descartes: Volume 1*. Cambridge: Cambridge University Press.

Descartes, R. 1991. *The Philosophical Writings of Descartes, Vol. 3: Correspondence,* trans. by J. G. Cottingham, R. Stoothof, D. Murdoch, and A. Kenny. Cambridge: Cambridge University Press.

Gabbay, D. M.; Guenthner, F. (eds.). 1983. *Handbook of Philosophical Logic*. Dordrecht, Netherland: Kluwer Academic Publishers.

Kripke, S. A. 1980. *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*. Cambridge, MA: Harvard University Press.

Lovejoy, A. O. 1936. *The Great Chain of Being: A Study of the History of an Idea*. Transaction Publishers.

Mill, J. S. 1843. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*. London, England: Longmans, Green, Reader, and Dyer.

Nolan, L. 2014. The Third Meditation: causal arguments for God's existence. In: D. Cunning (ed.) *The Cambridge Companion to Descartes' Meditations*. Cambridge University Press.

Notes

[1]"It does not make much difference whether my second proof ...is regarded as different from the first proof, or merely as an explanation of it [...] Nevertheless, it seems to me that all these proofs based on [God's] effects are reducible to a single one" *To [Mesland], 2 May 1644* (Descartes, 1991).

[2]For sake of brevity in the formulation of the argument I will consider just substances and modes.